

“Het datawarehouse wordt steeds meer gezien als een asset”

KLM vliegt voor de formatie uit



Paul van der Linden

KLM heeft een datawarehouse waarin niet alleen gegevens ‘rusten’, maar dat ook operationele activiteiten ondersteunt. De luchtvaartmaatschappij doet een aantal zaken anders dan wat in Nederland usance lijkt. Hier vraagt men zich af hoe het warehouse een rol kan spelen bij een fusie - niet ondenkbaar in dit geval. Maar ook wordt gezocht naar de beste manier om een divisiebreed datawarehouse op te waarderen tot een systeem dat de totale onderneming kan behapen. “Wat wij nu managementinformatie noemen, kan overmorgen operationele informatie zijn geworden.”

De basis van zijn datawarehouse legde de KLM begin 1998, vertelt Anton Hazewinkel, informatiemanager voor architectuur en managementinformatie binnen de Passage Divisie (PD). In de loop van dat jaar raakte Hazewinkel betrokken bij het warehousetraject. De eerste OLAP-toepassingen dateren overigens al van begin jaren negentig. Deze werkten op de voorloper van het datawarehouse, dat men bij KLM *information warehouse* (IW) noemt. Momenteel wordt het IW geleidelijk afgebouwd ten gunste van het huidige datawarehouse. Het platform is in al die jaren hetzelfde gebleven: MVS (OS/390). Zestig tot zeventig datamarts stammen uit die tijd en zijn, aldus Hazewinkel, niet meer weg te denken. “Daarmee hebben we redelijk voorop gelopen. We liepen

geregeld aan tegen de grenzen van de toenmalige technologie. Het accent lag met name op de verwerkingscapaciteit binnen de batchwindows.” Terugkijkend

“Vanuit het besturingsmodel moet de managementinformatie top-down en eenduidig worden geïmplementeerd”

beoordeelt hij de datawarehouse-industrie van die dagen als tamelijk onvolwassen.

Het Passage-datawarehouse wordt gebruikt voor Sales, Revenue Management en Marketing. PD is een informatie-intensief bedrijf. Hierdoor ligt het voor de

hand dat het PD-datawarehouse aan het uitgroeien is tot een KLM-brede ‘hangar’ van gegevens, hoewel de bouw ervan is begonnen in een tijd dat de ICT-organisatie nog sterk was gedecentraliseerd. Hazewinkel: “Inmiddels is weer sprake van een duidelijke centralisatie. Dat is ook nodig: steeds meer initiatieven hebben een proces-overstijgend karakter. De behoefte aan afstemming en integratie -denk bijvoorbeeld aan de processen met klantcontacten- vertaalt zich naar het belang van een centraal opgezet datawarehouse.”

GELEIDELIJKE OVERGANG

Terugkomend op de oorsprong van het huidige warehouse geeft Hazewinkel aan

Het onmisbare datawarehouse...

In 1999 verscheen een zestal beschrijvingen in Database Magazine over datawarehouses in Nederland, bij de ondernemingen Brocacef, Glaxo Wellcome -inmiddels Glaxo Smith Kline genaamd-, Grolsch, Mattel, NAM en Quaker Chemical. Uitgangspunt voor deze reeks was na te gaan in hoeverre ‘juichende’ verhalen over de toepassingen van datawarehouses daadwerkelijk opgaan in de Nederlandse situatie. Conclusie uit deze zes cases: alle beschreven datawarehouses zijn inmiddels onmisbaar. Zij leveren een wezenlijke bijdrage aan het bedrijfsproces. Maar het gaat hier voornamelijk om ondersteunende warehouses, die inderdaad een wat ‘relaxter’ karakter hebben dan de transactionele systemen waarvan ze hun data ontvangen. De reden om aandacht te besteden aan het KLM-datawarehouse is drieledig. In de eerste plaats ondersteunt dit warehouse duidelijk ook operationele activiteiten. Ten tweede heeft de luchtvaartmaatschappij een aantal keuzen gemaakt die duidelijk afwijken van het gemiddelde op basis van de zes eerste cases (zie tabel 1). Ten slotte spelen hier vragen als: ‘hoe kan het datawarehouse een rol spelen bij overnamen of fusies?’ en ‘hoe kun je een divisiebreed datawarehouse opwaarderen tot een corporate-wide datawarehouse?’. Deze vragen geven aan dat KLM met zijn datawarehouse een aantal stappen vooruit loopt.

Verschillen	Toelichting
notenbalk	<ul style="list-style-type: none"> • gewaardeerd communicatiemiddel tussen business en IT • geeft per dimensie eenduidig aan op welke niveaus gegevens worden getoond
<ul style="list-style-type: none"> • datawarehouse • software/hardware/besturingssysteem 	<ul style="list-style-type: none"> • datawarehouse is een DB2-database op een MVS (OS/390-)platform (mainframe-gebaseerd) • een Unix-platform; geen Oracle-database
dimensioneel datawarehouse	geen genormaliseerd relationeel model, maar ster schema's met conformed dimensions ('de grote ster')
eigenaarschap van data	expliciet belegd als onderdeel van Enterprise Architecture Program, dat wordt ingevuld door samenwerking tussen CIO Office en Passage Divisie
meerdere toepassingen	<ul style="list-style-type: none"> • verschillende organisatie-onderdelen (Sales, Revenue Management en Marketing) • verschillende toepassingen (Inventory Management, Monitoring Business Results en Pricing) maken gebruik van de warehousedata
conceptueel modelleren (FCO-IM)	conceptueel modelleren gaat vooraf aan het logisch datamodelleren. Voor het conceptueel modelleren wordt gebruik gemaakt van FCO-IM (een NIAM-variant)

TABEL 1. WAARIN KLM'S WAREHOUSING AFWIJKT VAN ZES DB/M-CASES.

dat het IW nooit een echt datawarehouse is geweest. "Je moet het meer zien als een platform voor cleansing, transformatie en aggregatie." De overgang van IW naar DW voert langs de weg der geleidelijkheid. Als bronnen op het DW worden aan-

gesloten, maakt dat oude applicaties, die deze bronnen gebruikten via het IW, overbodig. Hazewinkel gaat ervan uit dat het toch nog wel zo'n drie jaar duurt voordat het IW helemaal is overgezet. Hierbij spelen uiteraard ook de benodigde vaardigheden mee, evenals de kosten die ermee gemoeid zijn.

COMPONENTEN

Het datawarehouse bestaat uit een DB/2-database op een MVS-platform. Dit wijkt af van wat de zes eerder besproken bedrijven (zie kader *Het onmisbare datawarehouse...*) hebben gekozen: Unix en WNT. Destijds is niet overwogen te kijken naar een ander platform dan het reeds aanwezige IBM-mainframe, dat werd geselecteerd op basis van de reeds aanwezige skills en voorzieningen. De data worden ontsloten voor gebruikers met behulp van Essbase (nu van Hyperion) en Comshare/Decision.

Bij KLM zijn grote hoeveelheden gegevens nodig om analyses te kunnen doen. Anton Hazewinkel benadrukt de beper-

kingen die multidimensionale databases hier opleggen. Voor bepaalde toepassingen wordt momenteel ook de beweging terug naar relationale databases gemaakt, zowel vanwege de grootte van de kubussen die kunnen worden gebouwd als met de opbouwtijd ervan. Beide factoren zijn bij KLM van groot belang. Ook het aantal gevraagde dimensies kan een rol spelen. Hazewinkel legt uit dat het de bedoeling is steeds te werken met kleinere kubussen, en dat steeds meer gegevens worden opgeslagen in relationele databases. Met andere woorden: KLM kijkt naar HOLAP-oplossingen.

Als ETL¹-tool is MetaSuite in huis gehaald. Het gebruik daarvan komt de overzichtelijkheid van het gehele proces ten goede, vindt Hazewinkel. Als nadeel geldt daarentegen de tijd die nodig is om de expertise op te bouwen. Het gebruik van MetaSuite is voorgeschreven, maar in het unload-proces kan eventueel SAS worden ingezet. Met MetaSuite wordt ervoor gezorgd dat alle transformatieregels in één tool aanwezig zijn. Voor de datamarts gebruikt men MS SQL Server onder Windows NT. Bij PD lopen momenteel twee pilotprojecten met Impromptu en PowerPlay als frontend op de SQL Server-datamarts.

METADATA EN DATAMODELLERING

Voor datamodellering maakt KLM gebruik van FCO-IM en ERwin. FCO-IM is een NIAM-variant, die wordt ondersteund met een tool van Ascaris². Deze levert een conceptueel datamodel op. Dit model wordt vervolgens automatisch omgezet naar een genormaliseerd logisch datamodel en een stermodel. Deze stukken worden afgedekt met ERwin (huidige leverancier: CA).

Het betreft hier een standaard aanpak binnen PD. FCO-IM wordt onder meer gebruikt als communicatiemiddel naar gebruikers. Anton: "Maar ook de data-architecten vinden FCO-IM erg prettig, omdat het een beter overzicht geeft." Daarbij speelt nog iets anders een rol. Door de complexiteit van de materie kun je het overzicht in een tool als ERwin

FOTO'S CHRIS VAN BEEK



ANTON HAZEWINKEL: "WE RICHTEN ONS NU VOORAL OP ANALYTISCHE PROCESSEN, MAAR VOORZIEN DAT GAANDEWEG STEEDS MEER ANALYSES EN BESLISSINGEN WORDEN GEAUTOMATISEERD."

allicht verliezen. VSS wordt gebruikt om de ERwin-modellen te beheren.

Beheer van de metadata vormt een extra aandachtspunt. KLM gebruikt verschillende tools voor de opslag van metadata. Hazewinkel: "Wij hebben nog geen tool gevonden waarin alle metadata zijn te beheren. En op conceptueel niveau heb-

architectuurprogramma, dat principes vaststelt en richtlijnen aanreikt voor de verschillende architectuurdomeinen, waaronder datamanagement.

ORGANISATIE

Gebruikers zijn inmiddels gewend geraakt aan het gebruik van de "notenbalk". Dit is een grafiek waarbij dimensies worden afgezet tegen de verschillende niveaus. De kruispunten vormen hierbij de measures. De specificatiekracht van de notenbalk wordt als bijzonder sterk ervaren. Hiermee is het een uitstekend communicatiemiddel geworden tussen gebruikers en IT.

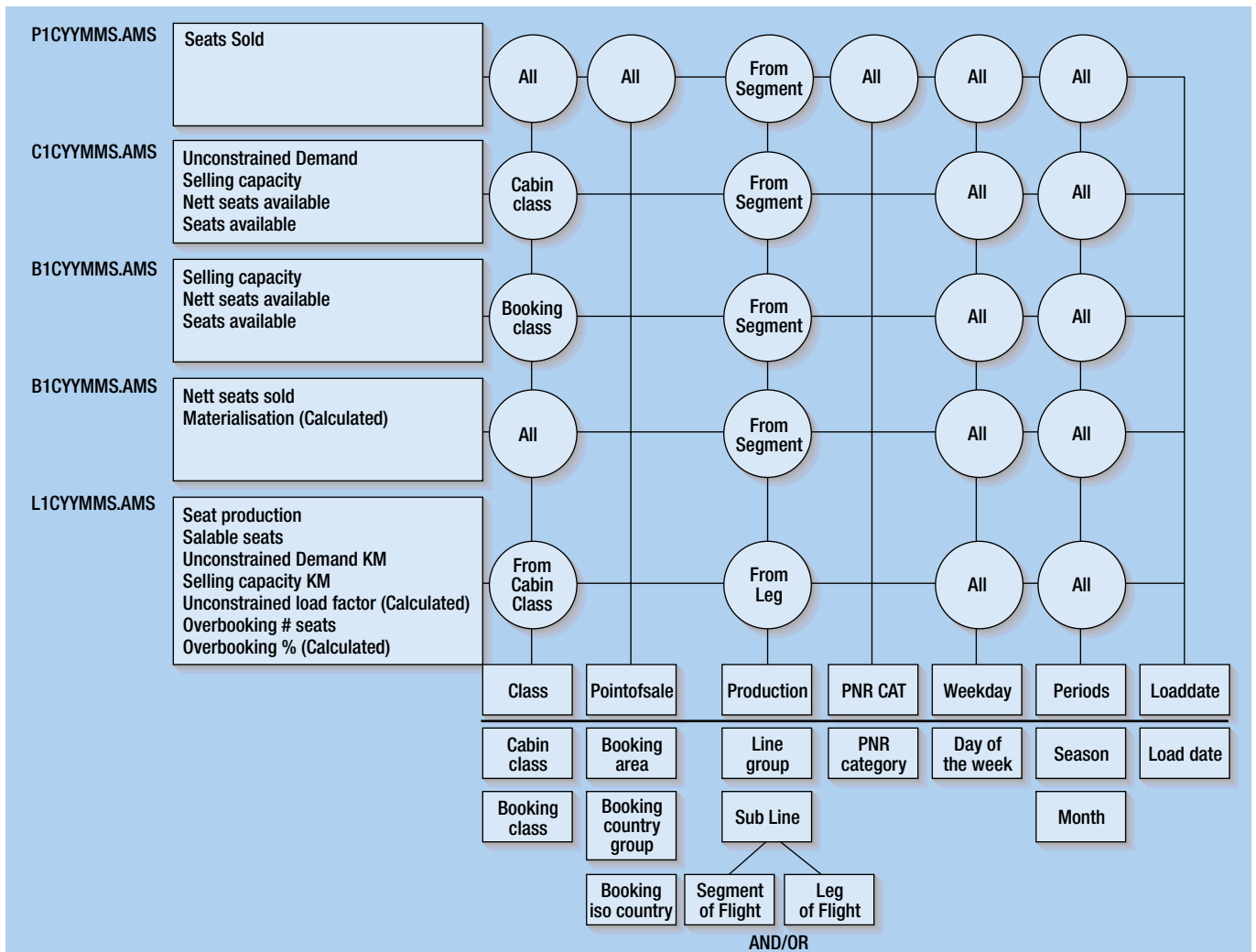
KLM's gebruikers zijn te vinden in de verschillende managementlagen en in de operationele processen. Men is nu bezig om vanuit de bestaande client/server-opstelling over te gaan naar webtoegang.



ANTON HAZEWINKEL: "ONS INFORMATION WAREHOUSE MOET JE MEER ZIEN ALS EEN PLATFORM VOOR CLEANSING, TRANSFORMATIE EN AGGREGATIE."

"Onderschatting ligt constant op de loer, vooral van wat wel of niet wordt aangetroffen bij de bron"

ben we nog te weinig discipline voor de vastlegging van metadata. Wat we het best op orde hebben, zijn de procesmodellen, de logische datamodellen en KPI-definities." Momenteel loopt een KLM-breed



FIGUUR 1: VOORBEELD VAN EEN NOTENBALK ZOALS DOOR KLM PASSAGE WORDT GEBRUIKT.

Er loopt een pilot met Cognos-tools. Wereldwijd hebben de verschillende KLM-vestigingen via het intranet toegang tot de datamarts. Deels wordt nog gebruik gemaakt van fat clients, maar dit zal verdwijnen op termijn. "Met het beschikbaar stellen van gegevens in het datawarehouse richten we ons nu vooral op analytische processen, maar we voorzien dat gaandeweg steeds meer analyses en beslissingen worden geautomatiseerd", vertelt Anton

Hazewinkel over de aangeboden data. "Wat wij nu managementinformatie noemen, kan overmorgen operationele informatie zijn geworden."

Het gebruik van datamarts wordt gemonitord. Bij een signalering dat bepaalde gebruikers weinig of geen gebruik maken van bepaalde data waartoe zij toegang hebben en waarvoor een licentie is afgesloten, wordt nagegaan wat de oorzaak hiervan is. Misschien zijn de

gegevens niet meer nodig; misschien is er een ander beletsel. Met de inrichting van een centrale business-afdeling, dus een niet-IT-afdeling, wil KLM dit nog veel breder en consequenter neerzetten. Elk kwartaal gaat een datawarehouse balanced scorecard naar het senior management.

De divisie-controller is de eigenaar van het proces voor managementinformatie. Die status bevordert de eenduidigheid binnen de divisie. "Het is belangrijk dat die ook terugkomt op alle niveaus van reporting en in alle rapportages", aldus Hazewinkel. "Vanuit het besturingsmodel moet de managementinformatie *top-down* en eenduidig worden geïmplementeerd."

ONTWIKKELING

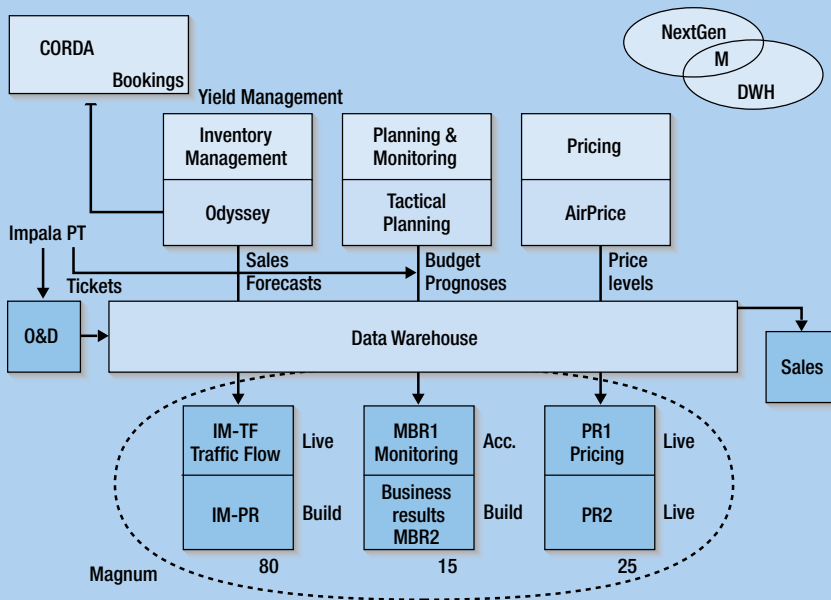
Anton Hazewinkel vervolgt: "Voor tweede derde laten we ons leiden door de vraag naar gegevens; de rest van de investeringen gaat naar centrale voorzieningen in het datawarehouse. Steeds vaker is de behoefte aan nieuwe informatie niet gereleerd aan één proces, maar aan meerdere processen. Het datawarehouse moet daarin voorzien. Hier worden immers alle procesgegevens verzameld en aan elkaar gerelateerd."

Het batchwindow is op dit moment nog geen probleem. Dit wordt ook niet verwacht, zolang het datawarehouse nog geen operationele taken heeft. Hierbij wordt gedacht aan toepassingen binnen customer relationship management (CRM). Momenteel heeft het warehouse voornamelijk analyse- en reportingtaken. In het ontwikkeltraject kan het wel voorkomen dat een toepassing wordt gevraagd waarvan van te voren duidelijk is dat die niet past binnen het batchwindow. Dan wordt de functionaliteit uit elkaar getrokken en ontstaan een aantal datamarts in plaats van één. De meeste gegevens worden wekelijks ververst; andere data kennen een maandelijkse frequentie. Dagelijks komt nog niet voor. "Dit zou een probleem worden, omdat je voor het opbouwen van sommige datamarts het weekend wel nodig hebt", zegt Hazewinkel.

Het gebruik van het datawarehouse

Magnum

Anthony Ithurbure is projectmanager voor de applicatie Magnum. Die maakt enerzijds deel uit van het datawarehouse-programma; anderzijds valt ze ook onder het programma Next Generation Revenue Management. Magnum bestaat uit een zestal modules die hun gegevens ontvangen uit het datawarehouse. Die modules bestrijken de aandachtsgebieden inventory management, planning & monitoring en pricing. Naast Magnum bestaan nog andere toepassingen op het warehouse, bijvoorbeeld ten behoeve van sales. Het datawarehouse zelf ontvangt gegevens vanuit O&D (yield management), Tactical Planning (planning en monitoring) en AirPrice (prijsinformatie). Daarnaast levert ook Impala (het ticketadministratiesysteem) informatie aan het datawarehouse aan. Het aantal directe gebruikers van Magnum bedraagt momenteel 120 personen



FIGUUR 2: OVERZICHT MAGNUM.

Het grote voordeel van Magnum, dat bestaat uit Essbase-kubussen met Decision en Excel als frontend, is dat op een redelijke intuïtieve wijze door de data genavigeerd kan worden. Doel is uiteraard een optimale bezetting van de vliegtuigen en een zo hoog mogelijk rendement. De verwachtingen voor bezettingsgraad en prijsniveau worden afgezet tegen de feitelijke gegevens. Een onderdeel hiervan betreft de beschikbaarheid van (vliegtuig)stoelen per verkooppunt voor bijvoorbeeld specifieke perioden en markten. Resultaten zijn hergroepeerbaar en leveren een direct antwoord op vragen als: waar is nog capaciteit over? Wat is de resulterende netto waarde? De uitkomsten hiervan kunnen aanleiding zijn tot het doen van acties, zoals discounts.

wordt niet doorbelast. De ontwikkeling ervan wordt bekostigd vanuit een centraal budget. In de toekomst wil men wel overgaan tot *unit pricing*. Hierbij is cruciaal dat het datawarehouse 'kritische massa' heeft bereikt. Dat betekent dat alle *key process*-gegevens zijn terug te vinden in het datawarehouse. Doorbelasting heeft dan alleen betrekking op de voorkant van het warehouse: het beschikbaar stellen van data; niet op het proces waarbij operationele bronnen gegevens doorsluizen naar het warehouse. Het KLM-datawarehouse lijkt zo langzamerhand de kritische massa te hebben bereikt.

Archiving en retrieval vinden nog niet plaats. "We hebben daarvoor nog wel tijd", verklaart Hazewinkel. "Het staat op de agenda. Je kunt niet alles in één keer aanpakken." Gegevens gaan momenteel in backup naar tape.

Wat Anton Hazewinkel opvalt, is de naïviteit die in sommige bedrijven leeft over de complexiteit van het bouwen en beheren van een datawarehouse. "Onderschatting ligt constant op de loer. Vooral van wat wel of niet wordt aangetroffen bij de bron. Ook heeft men het vaak al over een datawarehouse, terwijl slechts sprake is van een datamart. De complexiteit van bijvoorbeeld conformed dimensions speelt dan al helemaal geen rol." Een ander aspect is de verandering in theorievorming die plaatsvindt. Denk hierbij aan de opkomst -en



"STEEDS VAKER IS DE BEHOEFTE AAN NIEUWE INFORMATIE NIET GERELATEERD AAN ÉÉN PROCES, MAAR AAN MEERDERE PROCESSEN. HET DATAWAREHOUSE MOET DAARIN VOORZIEN."

ondergang?- van de multidimensionale database, genormaliseerde data versus sterschema's en de eisen ten aanzien van de respons die het datawarehouse moet kunnen leveren.

AANDACHTSPUNTEN

Wat zijn de problemen? Hazewinkel: "Allereerst de continuïteit in skills en expertise. Er is veel tijd nodig om mensen in te werken in datawarehousing. Het borgen van kennis en ervaring krijgt daarom extra aandacht." De complexiteit zit hierbij zowel in de datawarehouse-architectuur, als in de complexiteit en dynamiek van de

luchtvaart. Het heeft volgens Hazewinkel twee jaar geduurd om een kritische massa te bereiken. Het proces is er één van afwisselend gericht zijn op de business(vraag) en de architectuur. Ook de overgang van een sterk gedecentraliseerde

"Er is veel tijd nodig om mensen in te werken in datawarehousing. Borging van kennis en ervaring krijgt daarom extra aandacht"

organisatie naar meer centrale sturing speelt een rol. Dat is een langzaam proces, omdat hiervoor autonomie moet worden ingeleverd. "Maar het datawarehouse wordt steeds meer gezien als een asset." ●

Referenties

1. Software voor het geautomatiseerd extraheren, transformeren en laden van data.
2. FCO-IM staat voor 'volledig communicatie-georiënteerde informatiemodellering' (*fully communication oriented information modelling*). Een uitvoerig artikel over de toepassing van deze (Nederlandse) methode heeft gestaan in DB/M 8/2000 en 1/2001: Peter Alons, *Op weg naar beheerst redundant metadatabeheer (1 en 2)*.

Drs. P.F.H. van der Linden (P.vanderLinden@Synergetics.nl) is senior consultant bij Synergetics Management Information Company.

De KLM-organisatie

KLM behoeft nauwelijks introductie. Het bedrijf is sinds jaar en dag één van de favoriete werkgevers in Nederland. De maatschappij vervoerde in het boekjaar 1999/2000 meer dan 15,7 miljoen passagiers en 576.000 ton vracht. Hiertoe beschikt zij over een vloot van 127 vliegtuigen, die voor een groot deel vliegen in een gecombineerde vracht/passagiersconfiguratie. Per 31 maart 2001 had KLM meer dan 30.000 mensen in dienst, van wie 26.000 in Nederland. Het concern kent een indeling in divisies, onder meer de Passage Divisie (PD), Vracht/Cargo en Engineering & Maintenance. De luchtvaartsector is in beweging. Luchtvaartmaatschappijen moeten om te overleven strategisch interessante allianties aangaan die een wereldomspannende netwerk van hubs opleveren. Zo werkt KLM al sinds 1989 samen met het Amerikaanse Northwest Airlines. Een voorgenomen alliantie met het Italiaanse Alitalia en later met British Airways gingen niet door. Het betekende dat de reeds gestarte passage-joint venture KLM/Alitalia weer uit elkaar getrokken moest worden. Dit terwijl die al had geleid tot een volledige integratie van het management, de verkoop van tickets, de marketing en vijftig nieuwe codeshare-bestemmingen. Naast de alliantiepartners zijn er de netwerk- en routepartners, zoals Malaysia Airlines, Kenya Airways en het Noorse Braathens.